The Factorial Structure of Four Temperament Styles and Measurement Invariance Across Gender and Age Groups

Rowiński, Tomasz ; Cieciuch, Jan ; Oakland, Thomas

Abstract: The Polish Temperament Styles Questionnaire (PTSQ), derived from Student Style Questionnaire (SSQ) was developed to measure four bipolar temperament styles: extroverted versus introverted, practical versus imaginative, thinking versus feeling, and organized versus flexible. The study focuses on factorial validity and measurement invariance (configural, metric, and scalar) across gender and age groups using data from 1,022 students ages 8 to19. Confirmatory factor analysis (CFA) supports the four factor model, and multigroup confirmatory factor analysis (MGCFA) supports measurement invariance for both age and gender groups.

DOI: https://doi.org/10.1177/0734282913487045

The Factorial Structure of Four Temperament Styles and Measurement Invariance across Gender and Age Groups

Tomasz Rowiński

Cardinal Stefan Wyszyński University in Warsaw, Poland

Wóycickiego 1/3 bud. 14

01-938 Warsaw, Poland

t.rowinski@uksw.edu.pl

Jan Cieciuch

University of Finance and Management in Warsaw, Poland

Thomas Oakland

University of Florida, FL, USA

This research was supported by grant no. N N106 132439 in the years 2010–2012 of the Polish Ministry of Science and Higher Education awarded to the second author.

Abstract

The Polish Temperament Styles Questionnaire (PTSQ), derived from Student Style Questionnaire (Oakland, Glutting, & Horton, 1996) was developed to measure four bipolar temperament styles: extroverted versus introverted, practical versus imaginative, thinking versus feeling, and organized versus flexible. The study focuses on factorial validity and measurement invariance (configural, metric and scalar) across gender and age groups using data from 1022 students ages 8-19. Confirmatory factor analysis supports the four factor model, and multigroup confirmatory factor analysis supports measurement invariance for both age and gender groups.

Keywords

temperament, childhood, factorial validity, measurement invariance

Temperament refers to stylistic and relatively stable traits that subsume intrinsic tendencies to act and react in somewhat predictable ways to people, events, and other stimuli (Joyce, 2010). The Student Style Questionnaire (SSQ; Oakland, 2012; Oakland, Glutting & Horton 1996) is used to measure four bipolar temperament styles in childhood and adolescence based on the work of Jung as well as Meyers and Briggs: extroverted versus introverted, practical versus imaginative, thinking versus feeling and organized versus flexible styles. The four bipolar styles assess qualities that impact children's academic and social development and therefore are of potential importance when used in an educational context (Oakland et al., 1996).

Extraverted-introverted styles describe where people generally derive energy. Those with an extroverted style preference generally are energized by the outside world, especially by contact with people. They generally learn by talking, enjoy large groups, have many interests and friends, and respond quickly. In contrast, those with an introverted style preference generally are energized their inner world of ideas or mental life. They generally learn by reflecting and writing, prefer small groups or solitude, have few interests and close friends, and respond hesitantly and cautiously.

Practical-imaginative styles describe how people generally prefer to code information and envision their world. Those with a practical style preference generally attend to facts, objects, and details. They generally are realistic and pragmatic, understand things literally, enjoy sequential learning, and notice details. In contrast, those with an imaginative style preference generally view the world in terms of generalizations, possibilities, and insights. They generally are insightful, visionary, theory oriented, enjoy metaphor and symbols, learn by intuitive leaps, and focus on generalizations.

Thinking-feeling styles describe how people generally prefer to make decisions. Those with a thinking style preference generally use objective standards and strive for honesty and

fairness. They generally are analytical and quizzical, value logic over sentiment, display brief businesslike interactions; strive for fairness, and favor truth and justice. In contrast, those with a feeling style preference generally use personal standards and strive for harmony. They generally are trusting, sympathetic, seek harmony, value sentiment over logic, display tactful and friendly interactions, strive for harmony, and are compassionate.

Organized-flexible styles describe when people generally prefer to make decisions. Those with an organized style preference generally like to use lists, finalize decisions, and make decisions as soon as possible. They generally want to plan and rely on schedule, are dependable, keep their personal space neat, and enjoy predictable structure. In contrast, those with a flexible style preference generally like to delay decisions and keep their options open. They generally are flexible in their commitments, seek opportunity for play, tolerate disorder in their personal space,, and enjoy surprise and changes (Oakland et al., 1996).

Prior research used the SSQ to examine children's temperament development in 21 countries (Oakland, 2012), thus leading to cross-national comparisons. Howevcr, despite the instrument's somewhat widespread use, few studies have examined its factorial structure and measurement invariance (Benson, Oakland, & Shermis, 2009). The Polish Temperament Styles Questionnaire (PTSQ) is the Polish version scale based on SSQ's items. The current study focuses on the factorial validity of the four bipolar model as well measurement invariance of the model across gender and age group. Information on measurement invariance provides evidence of the degree an instrument measures the same latent dimension(s) in all groups (Raykov, Macoulides, & Li, 2012). This information may be especially important in research with children if the validity of data from self-report measures are questioned due to the children's young ages and measurement demands.

Method

Participants

Data were collected on 1022 students (51% female), ages 8 through 19 (M = 14.1, SD = 3.3), from both urban and rural areas in central and eastern Poland in school that agreed to participate in the study. The sample was divided into two age groups: younger one, ages 8-14 (n = 512), and older one, ages15-19 (n = 510). Children whose parents granted permission to participate in the study completed the PTSQ during their school lessons in about 15-20 minutes. Volunteer university students trained in its administration methods administered the scale consistent with the SSQ's directions.

Instrument

The PTSQ was derived from the Student Styles Questionnaire (SSQ). Each of its 69 forced-choice items has two alternatives that allow for an assessment of preferred behaviors associated with one of four bipolar styles: extroversion (E) or introversion (I), practical (P) or imaginative (M), thinking (T) or feeling (F), and organized (O) or flexible (L). The reliability of the instrument was assessed by the Index of Quality of Saris and Gallhofer (2007), based on data collected in this study: EI = .82, PM = .77, TF = .73, OL = .86. The index corresponds to the correlation between the latent variable and the observed variables.

Data Analysis

According to Benson et al. (2008) SSQ items were assigned into parcels and parcels were introduced into confirmatory factor analysis (CFA) for the entire group and separately for the two age and two gender groups. Test for measurement invariance was carried out using multigroup CFA. This method involves setting cross-group constraints and comparing more restricted models with less restricted models (Byrne, Shavelson, & Muthén 1989; Chen, 2007). There are three levels of measurement invariance (Vandenberg & Lance, 2000): (1) configural (all groups have the same pattern of factor loadings), (2) metric (the factor loadings are constrained to be equal across the compared groups), and (3) scalar (indicator intercepts are constrained to be equal across groups). Scalar invariance is required to compare construct

means across groups; its presence indicates that the scales were used in a similar way in each group.

Cut-off criteria suggested by Chen (2007) were used to determine whether the fit of more restrictive models deteriorated significantly. In samples larger than 300, the criteria for identifying a lack of metric invariance compared with configural invariance were a change greater than .01 in CFI, supplemented by a change greater than .015 in RMSEA or a change greater than .03 in SRMR. The criteria for identifying a lack of scalar invariance compared with the metric invariance model were a change larger than .01 in CFI, supplemented by a change larger than .015 in RMSEA or a change larger than .01 in SRMR.

Results and Discussion

The tested model in AMOS 20 is presented on the Figure 1. Descriptive statistics and covariance matrix are available from the first author upon request. Parcels within each dimension, items within parcels, and loadings in CFA in all analyzed groups are presented in Table 1. The model fit coefficients for each group are reported in Table 2.

Put Figure 1 about here

Put Table 1 about here

Put Table 2 about here

The models for the total group, the two age groups, as well as for male and female meet standard criteria (e.g., Hu & Bentler 1999). Data on the global fit measures from the multigroup CFA for the two age groups and two genders are reported in Table 3.

Put the Table 3 here

Configural metric and scalar invariance were supported for both gender groups. Metric invariance also was supported for the two age groups. However, changes in CFI and RMSEA exceeded criteria needed to accept the most restrictive model indicating full scalar invariance. Partial scalar invariance was established after releasing intercepts in three parcels (parcel 1

loading on extraversion-introversion, parcel 2 on practical-imaginative, and parcel 3 on organized-flexible styles). Partial invariance is sufficient for conducting meaningful between-group comparisons (Steenkamp & Baumgartner, 1998). Therefore, comparisons of means between gender and age groups, while treating temperament in its dimensional form, is meaningful and justified because scalar measurement invariance was established.

To summarize, results from the CFA provide support for the test's theoretical four factor structure. Support for measurement invariance was found for both age and gender groups. Full metric invariance, supported across age and gender group, enables comparisons of temperament correlates. Also, full scalar invariance, evident across gender, and partial scalar invariance evident across age groups, enable comparisons of means across age and gender groups (Vandenberg & Lance, 2000). Establishing measurement invariance is especially important in cross-sectional studies with young children because of doubts whether the measurement instrument provides reliable results due to age-related cognitive differences yet similar demands within the measurement instrument. The results suggest the PTSQ items are understandable to children as young as age 8 and the structure of temperament is similar for younger and older children. Information from the four bipolar styles is thought to be useful in educational context because it provides insights on how children learn and how temperament-related instructional methods may facilitate learning, taking into account individual differences between children in their temperament styles (Oakland et al. 1996). The factorial structure and measurement invariance provide support for the meaningful use of temperament data from the PTSQ in research and practice.

References

Benson, N., Oakland, T., Shermis, M. (2009). Cross-National Invariance of Children's Temperament. *Journal of Psychoeducational Assessment*, *27*(1), 3-16.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*(3), 456-466.

Chen, F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464-504.

Hu, L., & Bentler, P. M. (1999). Cut off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1-55.

Joyce. D. (2010). *Essentials of temperament assessment*. Hoboken, NJ: Wiley

Oakland, T. (2012). Children's temperament in 21 countries. The 8th Conference of International Test Commission, Amsterdam.

Oakland, T., Glutting, J. and Horton, C. (1996). *Student Styles Questionnaire*. San Antonio, TX: The Psychological Corporation.

Raykov, T., Marcoulides, G., A., & Li, C-H. (2012) Measurement invariance for latent constructs in multiple populations: a critical view and refocus. *Educational and Psychological Measurement, 72*(6), 954-974,

Saris, W. E., & Gallhofer, I. N. (2007). *Design, evaluation, and analysis of questionnaires for survey research*. Hoboken, NJ: Wiley.

Steenkamp, J.B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*, 78-90.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices and recommendations for organizational research. *Organizational Research Methods*, *3* (1), 4-69.

Figures and tables

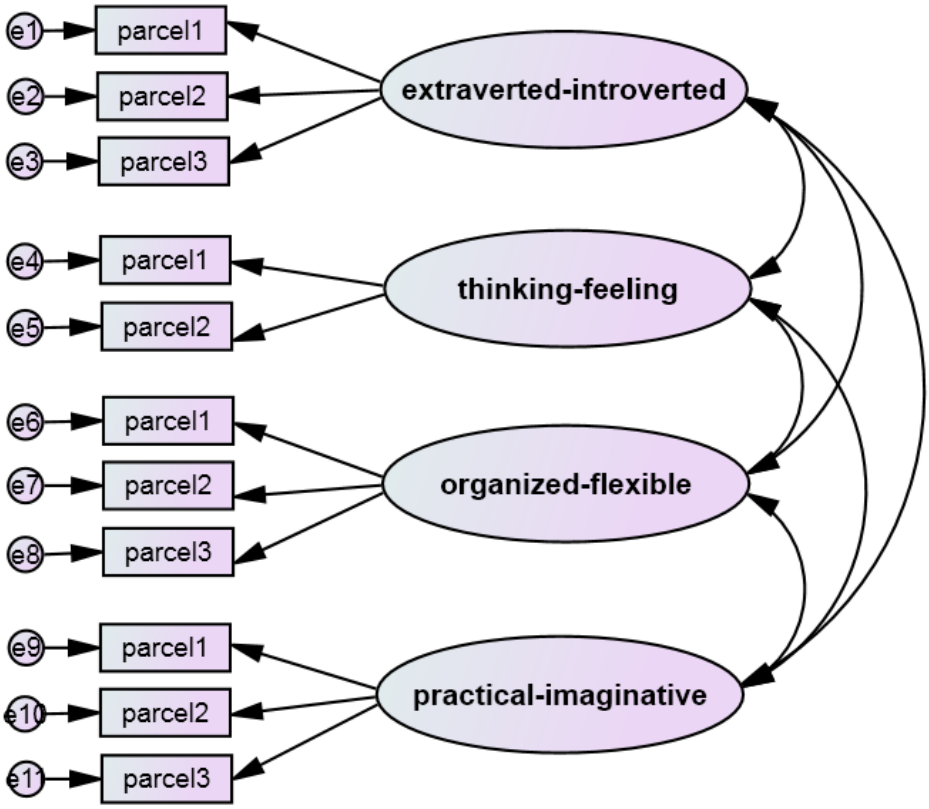

Figure 1. Measurement model of PTSQ tested in CFA.

Table 1

Parcels within each dimension, items within parcels and loadings in CFA in all analyzed groups

Dimension	Parcel	Items	Total group	Females	Males	Younger	Older
Extroverted -introverted	Parcel 1	1, 7, 19, 26, 46, 65	.49	.56	.41	.40	.65
	Parcel 2	4, 16, 36, 49, 62	.85	.78	.89	.89	.72
	Parcel 3	13, 39, 42, 52, 57, 67	.57	.60	.56	.58	.60
Practical-imaginative	Parcel 1	3, 11, 21, 45, 51	.55	.54	.48	.48	.57
	Parcel 2	6, 25, 34, 48, 64, 68	.55	.64	.36	.52	.57
	Parcel 3	9, 15, 29, 31, 41	.63	.64	.73	.68	.62
Organized-flexible	Parcel 1	14, 23, 32, 44, 53, 56, 60, 69	.71	.72	.71	.65	.68
	Parcel 2	5, 17, 27, 38, 40, 47, 59	.69	.75	.61	.61	.64
	Parcel 3	2, 8, 20, 35, 58, 63, 66	.71	.69	.72	.65	.70
Thinking-feeling	Parcel 1	18, 28, 43, 55	.68	.63	.79	.94	.75
	Parcel 2	12, 37, 50, 61	.53	.43	.41	.38	.50

Table 2

Global fit measures for the single sample CFAs with the PTSQ (df = 38)

	N	χ^2	*p*	CFI	RMSEA	SRMR
Female	520	119.2	<.001	.920	.064 (.051-.077)	.055
Male	504	105.7	<.001	.945	.060 (.046-.073)	.060
8-14 years old	512	109.4	<.001	.913	.061 (.048-.074)	.057
15-19 years old	510	62.2	<.001	.973	.035 (.018-.051)	.040
Total group	1022	167.7.0	<.001	.930	.058 (.049-.067)	.051

Table 3

Global fit measures for the MGCFA across gender and age groups with the SSQ

	χ^2	*df*	*p*	CFI	RMSEA	SRMR
Measurement invariance across gender groups						
Configural	224.9	76	<.001	.917	.044 (.037-.050)	.060
Metric	244.8	83	<.001	.910	.044 (.037-.050)	.063
Scalar	253.3	90	<.001	.909	.042 (.036-.048)	.063
Measurement invariance across age groups						
Configural	171.6	76	<.001	.944	.035 (.028-.042)	.057
Metric	192.7	83	<.001	.935	.036 (.029-.043)	.059
Scalar	299.6	90	<.001	.877	.048 (.042-.054)	.058
Partial scalar	208.6	87	<.001	.928	.037 (.031-.043)	.059